概率图模型笔记

宋佳欢

2019 年 10 月 10 日

目录

1 绪论

概率图的三个方面：

1. 表示：有向图（贝叶斯网络），无向图（马尔科夫网络），高斯图

2. 推断：精确推断，近似推断：确定性近似（变分推断），随机近似：MCMC

3. 学习: 参数学习，图结构学习

1.1 基础概念

高维随机变量的概率：

$$P(x_1, x_2, ..., x_p)$$

边缘概率：

$$P(x_i)$$

条件概率：

$$P(x_j|x_i)$$

加法法则（计算边缘概率）：

$$P(x_i) = \int P(x_1, x_2)dx_2$$

乘法法则（计算联合概率）：

$$P(x_1, x_2) = P(x_1) \cdot P(x_2|x_1)$$

链式法则（乘法的推广）：

$$P(x_1, x_2, \cdots, x_p) = \prod_{i=1}^{p} P(x_i|x_1, x_2, \cdots, x_{i-1}, x_{i+1}, \cdots, x_p)$$

贝叶斯定理：

$$P(x_2|x_1) = \frac{P(x_1, x_2)}{P(x_1)} = \frac{P(x_1, x_2)}{\int P(x_1, x_2)dx_2} = \frac{P(x_2)P(x_1|x_2)}{\int P(x_2)P(x_1|x_2)dx_2}$$

1.2 条件独立性

困境： $P(x_1, x_2, ..., x_p)$ 计算复杂，所以要简化：

1. 假设各个变量之间相互独立 (朴素贝叶斯)：$P(x|y) = \prod_{i=1}^{p} P(x_i|y)$

2. 现实性每个变量之间多少是有关联的，所以条件再放松一点，那就是马尔可夫性，

$$x_j \perp x_{i+1}|x_i, \quad j < i$$

3. 再推广，**条件独立性**假设:f

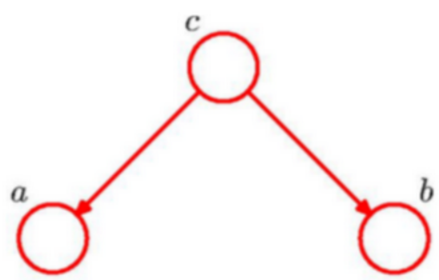
$$x_A \perp x_B|x_C, \quad x_A, x_B, x_C\text{是集合，且不相交}$$

条件独立性使用图来表示，在图上赋予概率的意义，使得图能表达条件独立性。

2 贝叶斯网络

2.1 贝叶斯网络的三种结构的独立性

1.tail to tail



计算三个变量的联合概率：

$$\text{因子分解：} \quad P(a, b, c) = P(c)P(a|c)P(b|c)$$

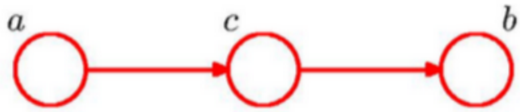
$$\text{链式法则：} \quad P(a, b, c) = P(c)P(a|c)P(b|a, c)$$

可得：

$$P(b|c) = P(b|a, c) \Longrightarrow a \perp b|c$$

该图的条件独立性：**若 c 被观测，则路径阻塞**（图论的说法，阻塞意味着独立）

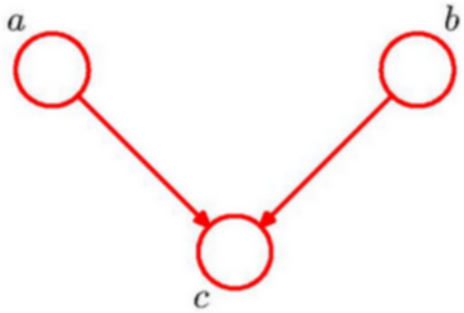
2.head to tail



f

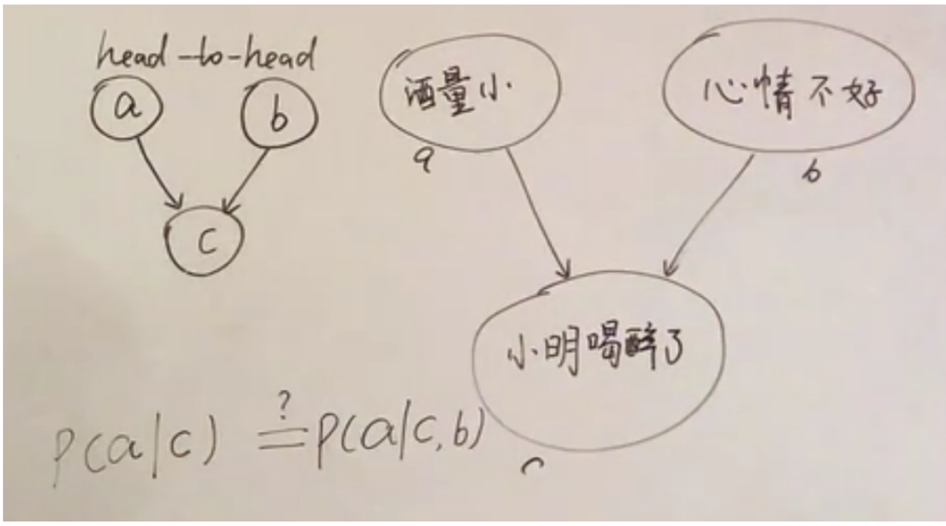
若 c 被观测，则路径阻塞,$a \perp b|c$

3.head to head



若 c 被观测，则路径是通的，即 a，b 之间不独立

一个例子：



得知小明喝醉，推小明酒量小的概率 $P(a|c)$ 应该是比较大的。

得知小明喝醉且心情不好，推小明酒量小的概率 $P(a|c,b)$ 就比上一种情况小了。因此 a，b 是相关的。

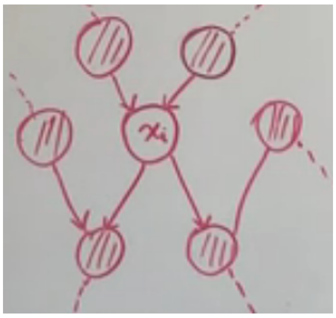
2.2 D 划分（d-Separation）

1. 集合 A,B,C，通过 tail to tail 和 head to tail 结构连接，A 到 B 的路径上的节点都必须在 C 的内部。

2. 集合 A,B,C，通过 head to head 结构连接，A 到 B 的路径上的节点以及其后继节点都<u>不能</u>在 C 的内部。

满足上述两个要求，则 $A \perp B|C$。即满足全局马尔可夫性。

以下图为例，计算边缘概率 $P(x_i|x_1, x_2, \cdots, x_{i-1}, x_{i+1}, \cdots, x_p)$:



$$
\begin{aligned}
P(x_i|x_1, x_2, \cdots, x_{i-1}, x_{i+1}, \cdots, x_p) &= \frac{P(X)}{P(x_1, x_2, \cdots, x_{i-1}, x_{i+1}, \cdots, x_p)} \\
&= \frac{P(X)}{\int P(X) dx_i} \\
&= \frac{\prod_{j=1}^{p} P(x_j|x_{parent(j)})}{\int \prod_{j=1}^{p} P(x_j|x_{parent(j)}) dx_i} \\
& x_{parent(j)}\text{表示}x_i\text{的父节点们}
\end{aligned}
$$

可见 x_i 的边缘概率只与和它相关的一些节点有关，即上图中的打阴影的节点,x_i 周围的一圈又叫马尔可夫毯。

3 Markov 网络

全局马尔可夫性：从节点集 A 中的节点到 B 中的节点的路径，都经过节点集 C 中的节点，则 $A \perp B|C$。

局部马尔可夫性：$a \perp \{全集 - a - 邻居\}|邻居$。

成对马尔可夫性：$x_i \perp x_j|\{全集 - x_i - x_j\}$

团: 图中节点的一个子集，其中任意两个节点有互相连接。

极大团：在极大团中再加入一个节点就不够成团。

因子分解 (极大团的势函数相乘)：

$$P(X) = \frac{1}{Z} \prod_{i=1}^{k} \psi(x_{C_i})$$